Fuzzy Music Query Retrievals from the Web Using Mobile Agents

Ali Selamat and Lim Sze Wei

Department of Software Engineering,
Faculty of Computer Science & Information Systems,
Universiti Teknologi Malaysia, 81310 UTM, Skudai, Johor, Malaysia.
Tel: +6-07-5532099; Fax: +6-07-5565044, Email: aselamat@fsksm.utm.my

Abstract. Internet is one of the most popular technologies that allow users to retrieve information from the Web. However, with the increasing number of web sites and music information have resulted poor performances in retrieving the relevant information. Non-formal English language web pages has also caused difficulties by the users in retrieving the relevance information from the Web. In this paper, we propose a Fuzzy Music Query Retrieval (FMQR) system based on mobile agent technology to support users in retrieving relevant music information form the Web. The main purpose of using mobile agent is to reduce network traffic. Whereas, the fuzzy sets theory is used to apply in fuzzy matching the music files that have higher degree of relevance to be retrieved by the system. From the experiments, the effectiveness of the fuzzy sets theory has been able to increase the degree of relevance on the music files retrieved by the mobile agents to be presented to the users.

1 Introduction

The Internet has become one of the main sources of entertainment, information retrieval and other purposes. There has been tremendous interest in the past few years in using mobile agent technology for automating the task of retrieving, organizing and filtering information located in the distributed databases across the World Wide Web (WWW) [1]. With the dramatic increasing on the number of documents and servers on the WWW, it is hard to quickly locate documents that contain potentially useful information. As a result, discovering and exploring the range of useful information source is an important problem to be solved. Furthermore, it is necessary to apply intelligent technologies to the problem of automatically analyzing Internet information sources [2] due to vast amount of information exist in the Web.

Research on fuzzy information retrieval using agent technology is one of the potential solution in order to overcome the problem. Delgado et al. [14] have presented a distributed intelligent agent model where the communication of the evaluation of the retrieved information among the agents is carried out by using linguistic operators based on the 2-tuple fuzzy linguistic representation as a way to endow the retrieval process with a higher flexibility, uniformity and

precision. The 2-tuple fuzzy linguistic representation model allows in making processes of computing with words without loss of information. Furthermore, Herrera-Viedma et al. [12] have presented a fuzzy linguistic multi-agent model that incorporates information filtering techniques in its structure, i.e., a collaborative filtering agent.The information filtering possibilities of multi-agent system on the Web are increased and its retrieval results are improved. However, the music information retrieval for a distributed web servers have not been discussed by the authors.

Other researcher, Vrettos [11] has discussed an intelligent agent for information retrieval and information filtering in the context of e-learning based on a fuzzy inference system which incorporates user profile knowledge. Also, Ko et al. [13] have proposed the searching mechanism based on ontology using mobile agent which is a platform independent and uses intelligent classifier among agents. This mechanism improves the efficiency of cooperation and information processing so that the user who requests for the location of information can find it accurately and rapidly. Moreover, Kim et al. [15] have proposed a conceptual information extraction from link-based search engine using a fuzzy concept network which can be personalized using the user's profile information. By combining personal fuzzy information retrieval and link-based search, the search agent provides high-quality information on the WWW about user query. However, the mobile agent performance have not been discussed in their experiments.

Therefore, in this paper we propose a Fuzzy Music Query Retrieval (FMQR) system which retrieved music information from the WWW using the mobile agents technology. The main purpose of using mobile agent is to reduce network traffic. Whereas fuzzy sets theory is used to apply in fuzzy matching, which music files that have higher degree of relevance will be retrieved and presented to the user. In the FMQR system, a keyword search process has been used for retrieving useful music information by combining a Porter stemming algorithm [3] and fuzzy sets theory [4-5]. The Porter stemming algorithm is use to strip off symbols and words ending by reducing them to a common core or stem word [8-9]. The fuzzy sets theory has been applied in fuzzy matching to retrieve music information that has higher degree of relevance [10]. From the experiments, the retrieval of music information by mobile agents has significantly reduce the time and disk space, because the agents have browsed through the network and perform the search locally on the remote server. Then the agents will only transfer useful music information, which has higher degree of relevance to user query [7].

The paper is organized as follows: The music information retrieval using fuzzy set theory is described in Section 2. The architecture of fuzzy music information retrieval using mobile agent is described in Section 3. The experiments and results are described in Section 4. The conclusion is described in Section 5.

2 Music Information Retrieval Using Fuzzy Set

According to Rijsbergen [3], information retrieval (IR) system has three main components: input, processor and output. We have created a model of the FMQR system based on the input from users query can be augmented by the words stored in a fuzzy dictionary. The augmented query will be sent to the FMQR system to be processed by the mobile agents. The results retrieved by the mobile agents will be given a feedback by the users in order to identify the relevancy of the query results. The framework for the music information retrieval process is shown in Fig. 1.

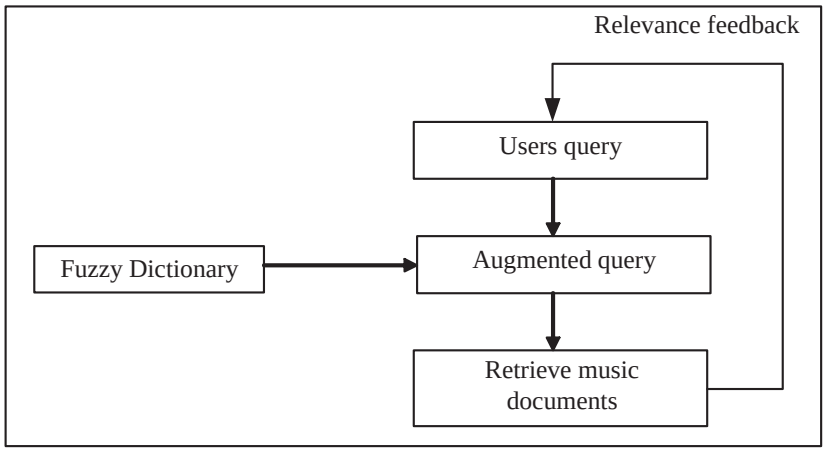

Fig. 1. The framework of music information retrieval process.

2.1 Fuzzy Music Query Retrieval

The problem of music information retrieval involve the process of identifying a set of word indexes (X) from a set of web pages (Y) where x_i is a set words that has been indexed. The set of words indexes (X) can be represented as follows:

$$X = \{x_1, x_2, x_3, .., x_i\}. \tag{1}$$

The related music web pages (Y) and y_i is a first web document and other related documents y_j, $j \simeq 2000$ can be represented as follows:

$$Y = \{y_1, y_2, y_3, .., y_j\}. \tag{2}$$

For the simplicity of the experiments we have limited $i, j \simeq 2000$ in order to test the efficiency of the retrieval system. In order to retrieve the fuzzy information, R , the formulation is given as follows

$$R : X \times Y \longrightarrow [0, 1] \tag{3}$$

where the membership value $R(x, y)$ show the relationships between words x_i and web document y_i.

As described in eq. 1, X is a classical set of words in the collection of web pages. The membership web pages in a predefined category A is often viewed as a characteristic function μ_A from X to $\{0, 1\}$ such that

$$\mu_A(X) = \begin{cases} 1 \text{ if } x \in A; \\ 0 \text{ if } x \ni A. \end{cases}$$

where $\{0, 1\}$ is called a valuation set; 1 indicates membership value while 0 indicates non-membership. If the valuation set is allowed to be in the real interval [0,1] , then A is called a fuzzy set. μ_A is the grade of membership of x in A. As the value of μ_A is closer to 1, the value of x is near to belong to A. Then, A can be represented as follows:

$$A = \{(x, \mu_A(X)), x \in A\}. \tag{4}$$

Also, the retrieval system has faced a problem in retrieving the synonym of the words. Fuzzy dictionary has been used to identify the synonym words with the query words. A query from a user will be represented as index of X words. Given U as a fuzzy set representing a query, the original query will composite with the fuzzy dictionary T which will produce a new augmented query B. The new augmented query is given by

$$U \circ T = B \tag{5}$$

where $\circ$ is the composite maximum-minimum. The pair value of composite

maximum-minimum is given by

$$B(x_j) = \texttt{maxmin}\lfloor U(x_i), T(x_i, x_j)\rfloor, \quad x_i, x_j \in X. \tag{6}$$

Therefore, from eqs. (3) and (6), for the new retrieved document D, the new augmented fuzzy keywords (B) used by the mobile agents to retrieve the music results based on the fuzzy set of keywords (Y) in the FMQR system which related to fuzzy information (R) can be formulated as follows:

$$B \circ R = D. \tag{7}$$

3 The architecture of fuzzy music information retrieval using mobile agent

In this section, a brief description on mobile agent technology and its relation to music information will be discussed.

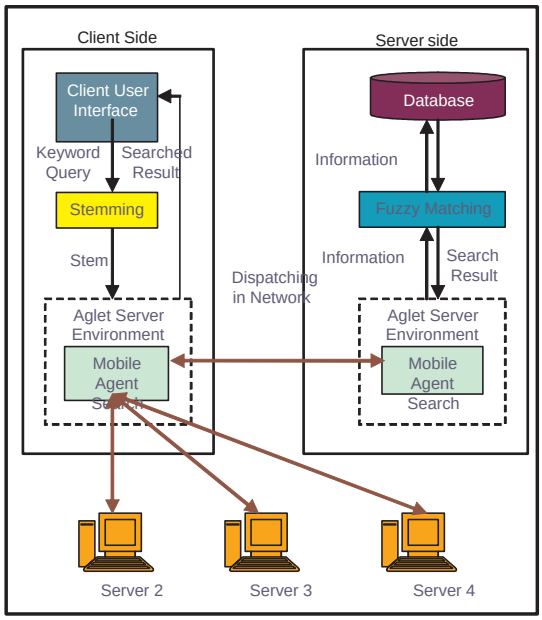

Fig. 2. The architecture of Fuzzy Music Query Retrieval (FMQR) system using mobile agent

3.1 Mobile agents

The introduction of mobile agent in the field of distributed computing has proven useful where the mobile agents will roam the networks to search for information requested by the users. The mobile agents will also cooperate with other agents to accomplish the assigned tasks. There are four main properties belonging to mobile agents such as intelligence, communication, autonomy, and mobility [1]. The biological insects have inspired most of the research related to agent based network routing and their colonies [2]. It relies on the principles that individual

insects will perform a simple behavior while the collective communities of these insects will perform complex problem solving capabilities [3]. A research has been conducted in mapping the biological insects to the network routing management by using mobile agents. These agents are represented as artificial agents that traverse the network to collect specific information from the designated hosts [5-6]. They will visit these hosts and coordinate with other agents to accomplish the assign tasks on behalf of users. They will also make several decisions to adapt their behavior according to the current environment in which they are currently resided.

3.2 Music Retrieval by Mobile Agents

Figure 2 shows the implementation of Fuzzy Music Query Retrieval (FMQR) system using mobile agent. At client side, user has to key in the search keyword which is either song tile or artist name in order to perform a search process. After the user key in the keywords, the system will perform a stemming process to the keywords by using Porter stemmer algorithms [3]. The new keyword created after the stemming process will be augmented with the relevance keywords from the fuzzy dictionary as shown in Fig. 1. The augmented keywords will then be passed to the mobile agent search to perform music information retrieval in the server that has previously been chosen such as server 2, server 3 and server 4. When the mobile agent search reaches to the server, it performs fuzzy matching with the web pages exist in the server database based on the formulae given in Eqs. (6) and (7). Then, the mobile agent will return the higher degree of relevance (DOR) of web pages together with the music files to the user at the client side.

As the web pages and music files will be returned to users at the client side, there is a chance for agents to retrieve the same file in different servers, resulting many redundant music files being retrieved at a time and increases the workload of mobile agents as well as the network traffic. In order to overcome this problem, the agents will identify the redundant music files and web pages and remove them from the lists before sending them to client.

The DOR on the music files retrieved by the mobile agents is formulated as follows:

$$\mathrm{DOR} = \frac{(\texttt{maxNumber} - \texttt{ResDiff})}{(\texttt{keywords} - \texttt{totalRec})} \tag{8}$$

where, the `maxNumber` is the maximum number of retrieved results, the `ResDiff` is the difference between the relevance results selected by the used compared to the results retrieved by mobile agents, the `keywords` is the number of keywords exist in the augmented query, and the `totalRec` is the total number of results which satisfied the user.

3.3 Fuzzy matching

When the mobile agent retrieve the query results, it will consider the accuracies of the augmented words and the stemmed words between the retrieved keywords and the music information. Also, the agent needs to interpret the following fuzzy inference rules as follows:

Rule 1 - IFK is **High**, **THEN** the search results is related.
Rule 2 - IF K is **Middle**, **THEN** the search results is consider related.
Rule 3 - IF K is **Low**, **THEN** the search results is not related.

where, K is the weight of the important degree of keywords match to database. The value **High** is between 0.7 to 1.0, **Middle** is between 0.5 to 0.69, and **Low** is between 0.0 to 0.29 as shown in Fig. 3. When the mobile agent selects the music information, which contain the high relevant information, the agent will retrieve the relevant information send return them to the client host. The user will se the results related to the weight of the music documents, title of the song, the artists name etc.

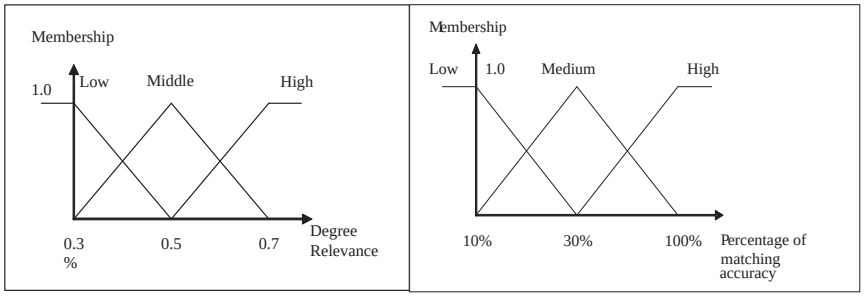

Fig. 5(a). Fuzzy set for the accuracy of query results with the augmented fuzzy keywords.

Fig. 5(b). Fuzzy set for the query weight which related with the degree of relevance on the music information.

Fig. 3. Fuzzy set for the accuracy of query results with the augmented fuzzy keywords.

4 The experiments and results

We have conducted two different experiments on the FMQR system. The user interface of the FMQR system is shown in Fig. 5. The results returned by the system is shown in Fig. 6. The first experiment is to validate the effectiveness of using the augmented fuzzy keywords for music query retrieval. The second experiment is to evaluate the time taken by mobile agents to retrieve the query results from different servers. For the first experiment, few examples on the

Fig. 4. The Fuzzy Music Query Retrieval (FMQR) system user interface

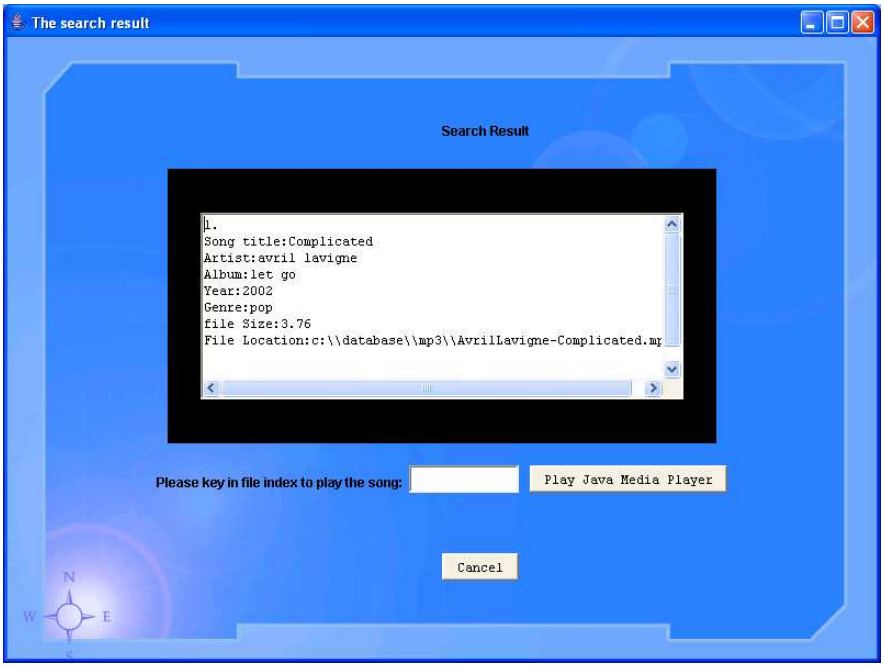

Fig. 5. The retrieved music query results

music files retrieved by using mobile agents from the selected servers are `avril lavgin`, `sacrifices` and `immortal`.

For experiment one, the results of retrieving different queries based on the fuzzy augmented keywords are shown in Figs. 7, 8, and 9, respectively. As a result, the degree of relevance (DOR) on the retrieval query results based on different types of music files has been increasing due to the support of fuzzy augmentation approach on the original keywords being retrieved by the system. For experiment two, the results of retrieving different queries based on the fuzzy augmented keywords are shown in Figs. 10 and 11, respectively. The time take to retrieve the query results from different servers can be minimized if the size of query results are small.

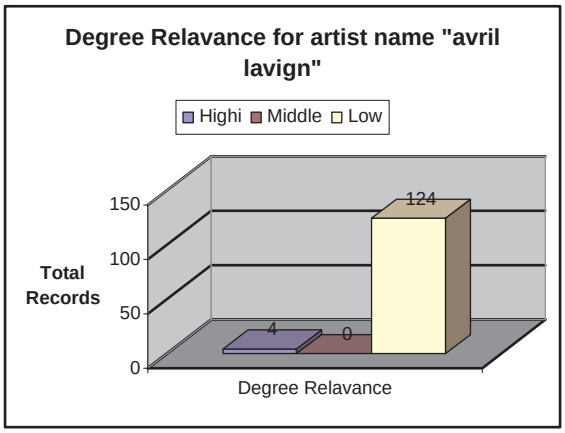

Fig. 6. Degree of relevance for song title `avril lavign`.

5 Conclusions

We have developed the Fuzzy Music Query Retrieval (FMQR) system in order to support Internet users for searching and retrieving the required information form the WWW. We have evaluated the system by given examples of queries, focusing on time information retrieval and the file size of information retrieval. The results shown in Fig. 5, are the local path of the files stored in the director from on of the servers. It shows the artist name, title, year, and the database filename. Further customization of the user interface for the FMQR system will be done in the future. However, the degree of relevance on the retieved results has been shown in Figs. 6, 7, and 8 respectively. From the conducted experiments, the results have shown that the increasing number of query results retrieved by the agents, the quality of search results will not be compromised. Further

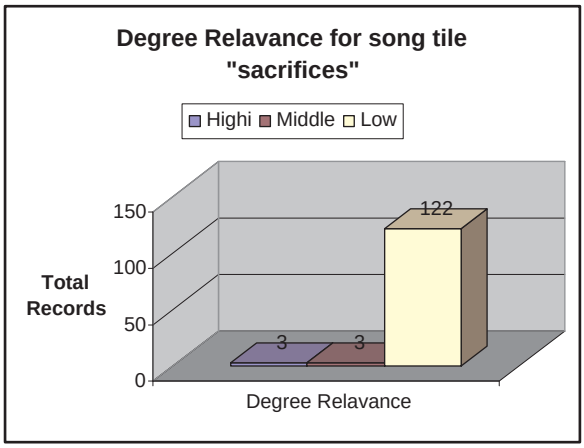

Fig. 7. Degree of relevance for song title `sacrifices`.

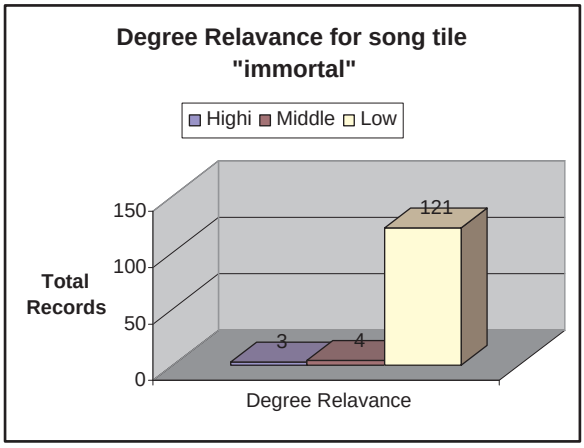

Fig. 8. Degree of relevance for song title `immortal`.

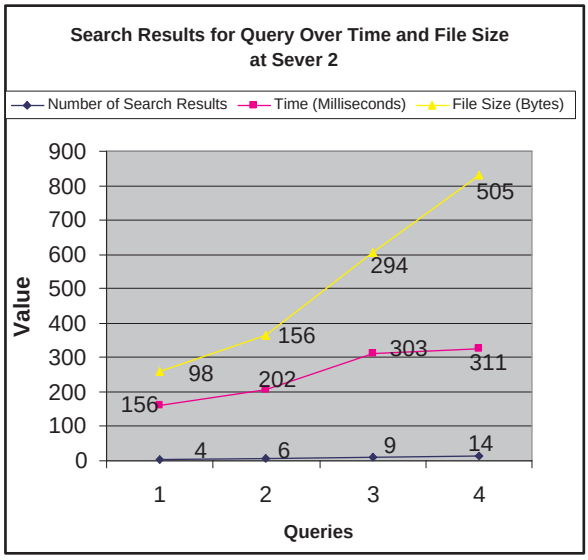

Fig. 9. Search results for query over time and file size at Server 2.

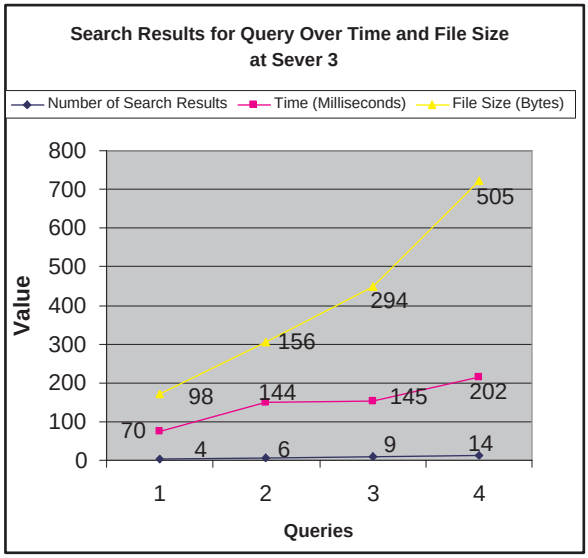

Fig. 10. Search results for query over time and file size at Server 3.

research on minimizing the time taken by mobile agents for retrieving a large number of query results from many servers is a subject under studies.

References

1. Selamat, A. and Selamat, M. H., Analysis on the Performance of Mobile Agents for Query Retrieval from Internet, Int. Information Sciences, Volume 172, Issues 3-4, June (2005) 281-307.
2. Glitho, R. H., Mobile Agents and Their Use for Information Retrieval: A Brief Overview and an Elaborate Case Study, Canada, IEEE Network, Jan./Feb. (2002) 34-41.
3. Rijsbergen, V. C. J., Information retrieval. 2nd ed. Butterworths, London. Department of Computing Science, University of Glasgow (2002)
4. Aliev, R.A. dan Aliev, R.R., Soft Computing and Its Applications. 1st ed. United States of America: World Scientific Publishing (2002)
5. Aglet Software Development Kit, http://sourceforge.net/projects/aglets/ (Accessed on January 12, 2004).
6. Ding Y-S. and Ren L-H., Merging Mobile Agents, Genetic Algorithms, and Fuzzy Logic for Intelligent Internet Search, IEEE, (2001) 811-816.
7. Jain, R. and Anjum, R, Mobile Agents for Personalized Information Retrieval: Why are they a good idea?, IEEE, (2000) 242-245.
8. Korfhage, R. R., Information Storage and Retrieval. 1st ed. United States of America: John Wiley & Sons (1997)
9. Loia, V., and Luongo, P., A Similarity-based view to Distributed Information Retrieval with Mobile Agents, IEEE, (2001) 1283-1286.

10. Chen, S-M, and Wang, J-Y., Document Retrieval Using Knowledge Based Fuzzy Information Retrieval Technique, IEEE, (1995) 793-803.
11. Vrettos, S.,and Stafylopatis, A., A Fuzzy Rule-Based Agent for Web Retrieval-Filtering, N. Zhong et al. (Eds.): WI 2001, LNAI 2198, pp. 448-453, 2001. Springer-Verlag Berlin Heidelberg 2001
12. Herrera-Viedma, E., Porcel, C., Gabriel Lopez, A., Olvera, A.D. and Anaya, K., A Fuzzy Linguistic Multi-agent Model for Information Gathering on the Web Based on Collaborative Filtering Techniques, J. Favela et al. (Eds.): AWIC 2004, LNAI 3034, pp. 3-12, 2004.
13. Ko, J., Gerardo, B.D., Lee, J., and Hwang, J-J. The Information Search System Using Neural Network and Fuzzy Clustering Based on Mobile Agent, Proceedings of Computational Science and Its Applications - ICCSA 2005: International Conference, Singapore, May 9-12, 2005, Part II, Editors: Osvaldo Gervasi, Marina L. Gavrilova, Vipin Kumar, et al. ISBN: 3-540-25861-2
14. Delgado, M., Herrera, F., Herrera-Viedma, E., Martin-Bautista, M. J., Martnez, L., Vila, M. A., A Communication Model Based on the 2-tuple Fuzzy Linguistic Representation for a Distributed Intelligent Agent System on Internet, Soft Computing - A Fusion of Foundations, Methodologies and Applications, Springer-Verlag GmbH, ISSN: 1432-7643,Volume 6, Number 5 , August 2002, Pages: 320 - 328
15. Kim, K-J., and Cho, S-B., Conceptual Information Extraction with Link-Based Search, Web Intelligence: Research and Development: First Asia-Pacific Conference, WI 2001, Maebashi City, Japan, October 23-26, 2001. Proceedings, ISSN: 0302-9743, Volume 2198 / 2001, Editors: N. Zhong, Y. Yao; Liu, S. Ohsuga (Eds.), p. 367, 2003.